Objective model for perceptual sharpness in image sequences

Gregory J. Power
Air Force Research Laboratory
Target Recognition Branch, AFRL/SNAT,
Building 620, 2241 Avionics Circle
Wright-Patterson AFB, Ohio 45433-7321

Kirk Sturtz
Veridian Incorporated
5200 Springfield Pike, Suite 200
Dayton, Ohio 45431

Abstract ***Currently, the analog video world is in transition to digital video. Along with that transition is the need for analysis techniques that can be applied to digital image sequences. In particular, objective techniques that can analyze the digital image sequences in a way that reflects human perception is being pursued. These objective measures can be valuable in many areas including quality assessment. Currently, subjective assessment of video is time consuming and expensive. Therefore, an objective assessment that accurately reflects the human visual system can be valuable. This paper examines the perceived quality of shifting and blurred sequences. It introduces an objective quality measure for assessing the sharpness in digital image sequences. The model is compared to two other objective measures and a subjective assessment.***

Keywords: Image Sequence Analysis, Quality, Objective Assessment, Data Storage and Retrieval, Motion Imagery, Digital Video

1 Introduction

The world of analog video is in quick transition to digital video which is also called digital image sequences. This transition is prompting researchers to investigate analysis techniques that can be used for digital image sequences. In particular, objective techniques that can analyze the digital image sequences in a way that reflects human perception is being pursued. These objective measures can be valuable in many areas including quality assessment. Currently, subjective assessment of video is time consuming and expensive. Therefore, an objective assessment that accurately reflects the human visual system or incorporates perceptual criteria can be valuable. Machine measures do not always match subjective measures. Comparison to a subjective measure can give credence to a machine measure. Care must be taken in the choice of video analysis techniques. One must not blindly use measures that were developed and successfully used for static image analysis to accomplish video analysis. Some psycho-physical experiments suggest that a different model is needed for motion imagery than is needed for static imagery. This paper introduces an objective quality measure for assessing the sharpness in digital image sequences. The model is compared to subjective assessments.

2 Perceptual Analysis

Perceptual research has demonstrated that looking at a static image is quite different than looking at an image sequence with object motion. This suggests that different analysis tools that mimic perceptual results are needed. Psycho-physical evidence [1] exists showing a trend that a sequence with a shifting subimage improves quality perception at low eye pursuit speeds unless the image sequence

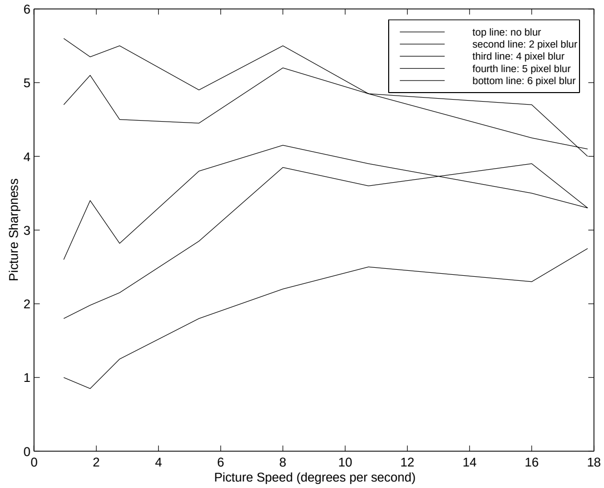

Figure 1: The graph shows the subjective results introduced by Westerink and Teunissen [1] that indicate that low quality images initially show increase in sharpness with an increase in speed, and high quality image sequences show a decrease in sharpness with increase in speed. The results are an average of 4 persons' observations and 12 repetitions of the data.

is already of high quality. Westerink and Teunissen [1] showed 4 subjects a moving portrait of a woman at various speeds and at various levels of linear blur. Four blurred sequences are obtained by applying 2, 4, 5, and 6 pixel linear blur functions to the original image sequence. Figure 1 shows the subjective results of this psycho-physical evidence. A similar experiment is recreated using a different portrait as shown in Figure 2.

2.1 $\hat{s}$ Objective Metric

An objective quality metric denoted as $\hat{s}$ was introduced by the Institute for Telecommunication Sciences (ITS) [2]. The metric was designed to work with digital image sequences suffering from compression artifacts. It was also designed to correlate with subjective assessments over image sequences that exhibit a wide variety of spatial and temporal informa-

Figure 2: This shows a single frame from each of the 5 image sequences used to test the objective metrics. In the experiment, the portrait rocks back and forth at various speeds in the rectangular black box. The top image is the original. The other frames are linearly blurred by 2, 4, 5, and 6 pixels.

tion content. The metric is defined as

$$\hat{s} = 4.77 - 0.922m_1 - 0.272m_2 - 0.356m_3. \quad (1)$$

To define the variables m_1, m_2, and m_3, some terminology needs to be defined. Using the ITS notation, the original sequence is defined as O_n and the lower quality sequence is defined as D_n. $SI[O_n]$ is the spatial information content based on the standard deviation of the edge information in the image, O_n. $TI[O_n]$ is the temporal information content based on the standard deviation calculated over the frame difference between successive images. Defining other symbols, STD is standard deviation, STD_{time} is standard deviation over time, MAX is maximum, MAX_{time} is maximum over time, RMS_{time} is root mean square over time, and $CONV$ is the convolution function. Now, the variables in $\hat{s}$ are defined as

$$m_1 =$$

$$RMS_{time}\left(5.81\left|\frac{SI[O_n]-SI[D_n]}{SI[O_n]}\right|\right), \quad (2)$$

$$m_2 =$$

$$f_{time}\left(0.108MAX\left\{(TI[O_n]-TI[D_n]),0\right\}\right), \quad (3)$$

$$m_3 = MAX_{time}\left\{4.23LOG_{10}\left(\frac{TI[D_n]}{TI[O_n]}\right),0\right\}, \quad (4)$$

where

$$f_{time} = STD_{time}\left\{CONV\left(x_i,[-1,2,-1]\right)\right\}. \quad (5)$$

$\hat{s}$ was designed to correlate well with subjective data sets that demonstrate all the digital artifacts introduced by lossy compression. However, as shown in Figure 3 with the tests on the moving portrait data set, $\hat{s}$ does not show the appropriate trends for a shifting portrait.

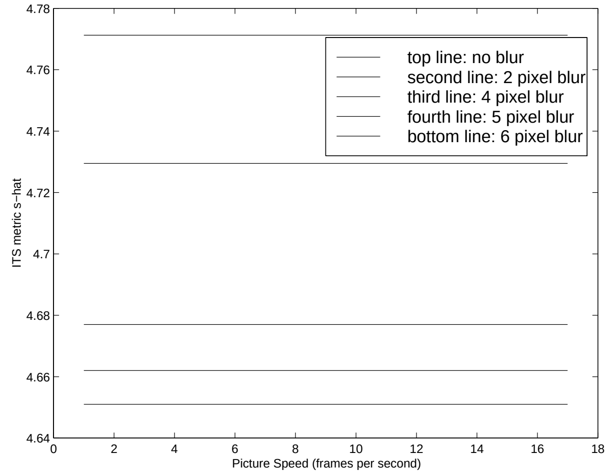

Figure 3: For image sequences of varying quality, $\hat{s}$ does not reflect psycho-physical evidence. $\hat{s}$ is constant for increases in the speed across the field of view.

2.2 Peak Signal-to-Noise Ratio (PSNR)

On occasion, the peak signal-to-noise ratio (PSNR) has been used to determine static image quality. Letting f denote an original image and g a linearly blurred image of the original. Let $\{f_k\}_{k=1}^{K}$ denote K copies of the original image f, and $\{g_k\}_{k=1}^{K}$ denote K copies of the degraded image g. We call f_k the k^{th} frame of the original image and similarly with g_k. Let $f_k(i,j)$ denote the image value at pixel position (i,j) for the k^{th} frame of the original image, where $i = 1,\ldots,M$ and $j = 1,\ldots,N$ range over all the pixels in the horizontal and vertical directions of the image. Similarly, let $g_k(i,j)$ denote the image value at pixel position (i,j) for the k^{th} frame of the degraded image. The error in the k^{th} frame at pixel position (i,j) is $e_k(i,j) = f_k(i,j) - g_k(i,j)$. The MSE at frame k, which is identical for all frames, is

$$MSE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} e_k(i,j)^2}{MN}. \quad (6)$$

and the PSNR is

$$PSNR = 20\log_{10}\frac{[\text{peak value of } f_k(i,j)]}{\sqrt{MSE}} \tag{7}$$

where the peak value of $f_k(i,j)$ is usually taken as 255 for an 8 bit image.

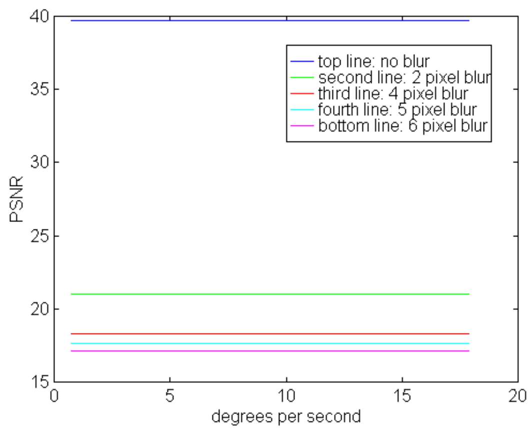

Figure 4: For image sequences of varying quality, the PSNR does not reflect psycho-physical evidence. PSNR only shows constant SNR for increases in the speed across the field of view.

Figure 4 shows PSNR does not show any improved quality. Therefore, a different and more robust metric is needed to analyze image sequence quality.

2.3 Modeling the Perceptual Increase in Sharpness

In this section, we introduce a model that shows itself to be a better quality metric than $PSNR$ and $\hat{s}$ in following the initial trends of subjective assessment. It was previously reported [3] that the increased sharpness with the initial increase in motion may be modeled by using an integration of image frames within the sequence. Using the notation developed in the last section, the mean error term, at pixel position (i,j), for K images is

$$\begin{aligned} \overline{e_k(i,j)} &= \tfrac{1}{K}\textstyle\sum_{k=1}^{K} e_k(i,j) \\ &= \tfrac{1}{K}\textstyle\sum_{k=1}^{K} (f_k(i,j) - g_k(i,j)). \end{aligned} \tag{8}$$

Under the hypothesis that the image is moving δ pixels per frame in the horizontal direction it follows

$$e_{k+1}(i,j) = e_k(i,j-\delta). \tag{9}$$

Given multiple frames, we define the MSE for integrated sequences (IMSE) as

$$IMSE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\frac{\sum_{k=1}^{K} e_k(i,j)}{K}\right)^2}{MN}. \tag{10}$$

Correspondingly, the PSNR for integrated sequences (IPSNR) is defined as

$$IPSNR = 20\log_{10}\frac{[\text{peak value of } f_k(i,j)]}{\sqrt{IMSE}}. \tag{11}$$

The following theorem shows that the IPSNR can increase in value as a portrait starts to shift/move linearly across a scene.

THEOREM. Let $\{f_k\}_{k=1}^2$ and $\{g_k\}_{k=1}^2$ be a family of frames of original and degraded images. Suppose
(1) the image is moving one pixel per frame in the horizontal direction, and
(2) the images satisfy $e_k(i,j) = 0$ for $j = 1, 2$ and $j = N-1, N$ for all i. ("Padded with zero's" or "black background")
Then

$$IMSE \leq MSE. \tag{12}$$

COROLLARY. Under the hypotheses of the theorem $IPSNR \geq PSNR$. (A one-dimensional proof can be found in reference [3]. An extended proof will be presented in a forthcoming publication.)

The results of calculating $IPSNR$ for the image sequence of Figure 2 are shown in Figure 5. As predicted by our theorem, $IPSNR$ shows an increase as the images begin to shift.

3 Conclusion

This paper suggests incorporating perceptual criteria to create objective quality measures. A test case is given which can be used to test objective measures against known subjective trends. Two existing objective measures are shown to fail to represent the trends for subjective assessment. If a certain quality level is needed, moving imagery is perceived at a different sharpness than a static image. This difference suggests some new measures are needed to take advantage of subjective perception for video.

A new criteria which integrates frames is introduced as an objective measure for determining perceptual sharpness. We then prove this new measure satisfies the qualitative condition which has been obtained experimentally. Hence, PSNR using integrated sequences is demonstrated as a technique that reflects some of the trends that are seen in subjective video perception. The challenge for the storage and retrieval of digital video is to closely consider the perceptual attributes of the motion imagery.

References

[1] Joyce HDM Westerink and Cornelis Teunissen. Perceived Sharpness in Moving Images. *Human Vision and Electronic Imaging: Models, Methods, and Applications.* Proc. SPIE, volume 1249, 1990.

[2] Digital Transport of One-Way Video Signals: Parameters for Objective Performance Assessment. *ANSI T1.801.03-1996*, February, 1996.

[3] Gregory J. Power. *Spatio-temporal feature-based objective quality metrics for digital image sequences.* Ph.D. Dissertation, University of Dayton, August, 1998.

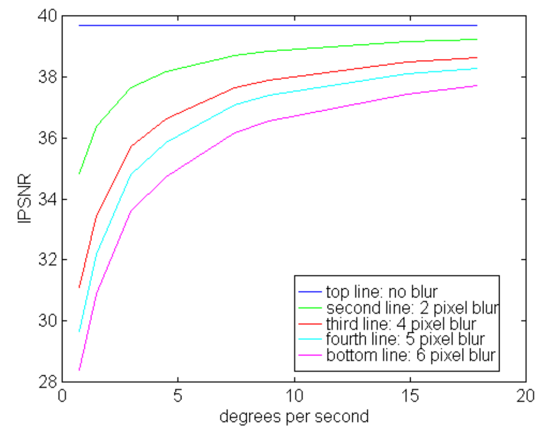

Figure 5: IPSNR agrees with psycho-physical evidence [1] showing the trend to increase SNR as degraded images initially increase in speed across the field of view.